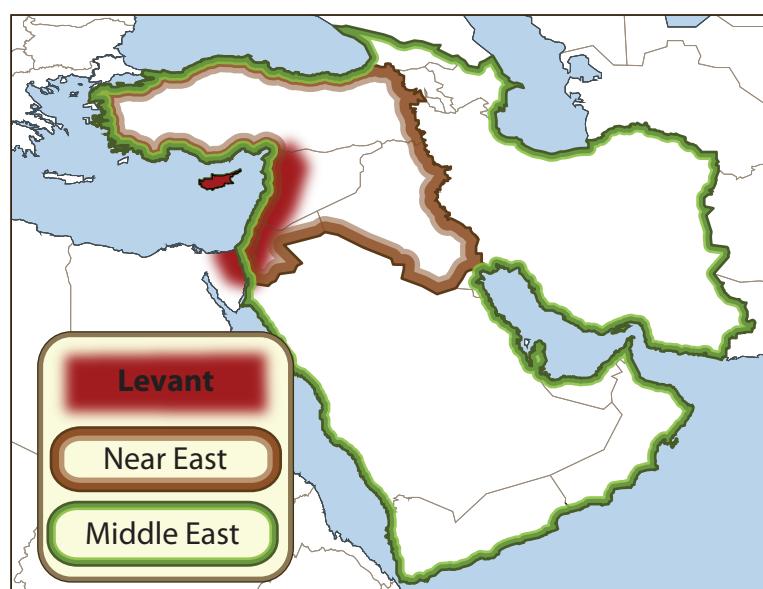


SUPPLEMENTARY INFORMATION

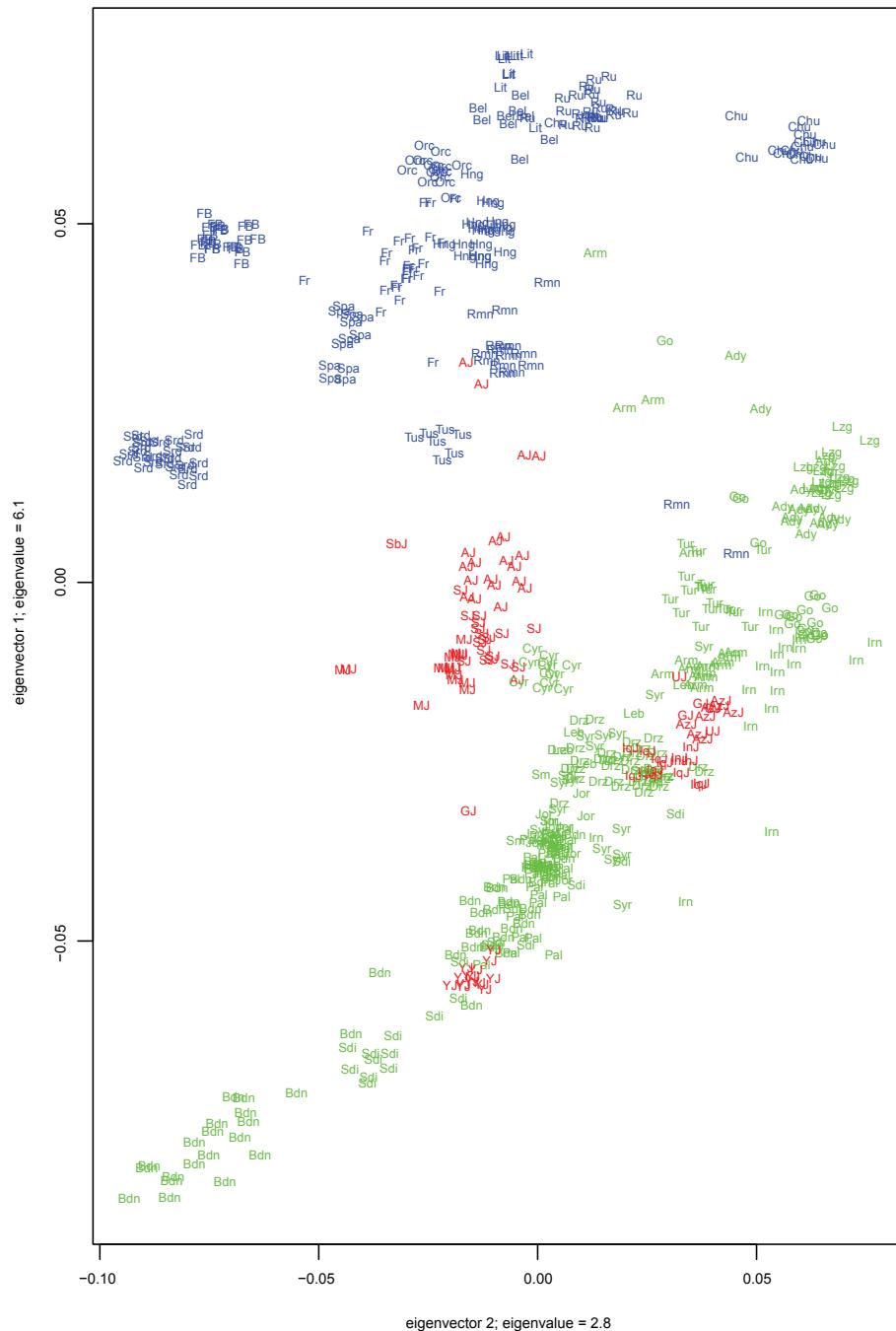


Supplementary Figure 1| Map of region names as defined in this study. We note that the terms Middle East, Near East, and Western Asia have overlapping and often inconsistent meanings, both historically and currently in different fields of study. We use the term Levant as a geographically- rather than politically- defined region. For simplicity, we limit our geographic terminology throughout the paper to the terms Middle East and Levant.

Figures 2a-c are in separate file - see 'Supplementary Figures 2'

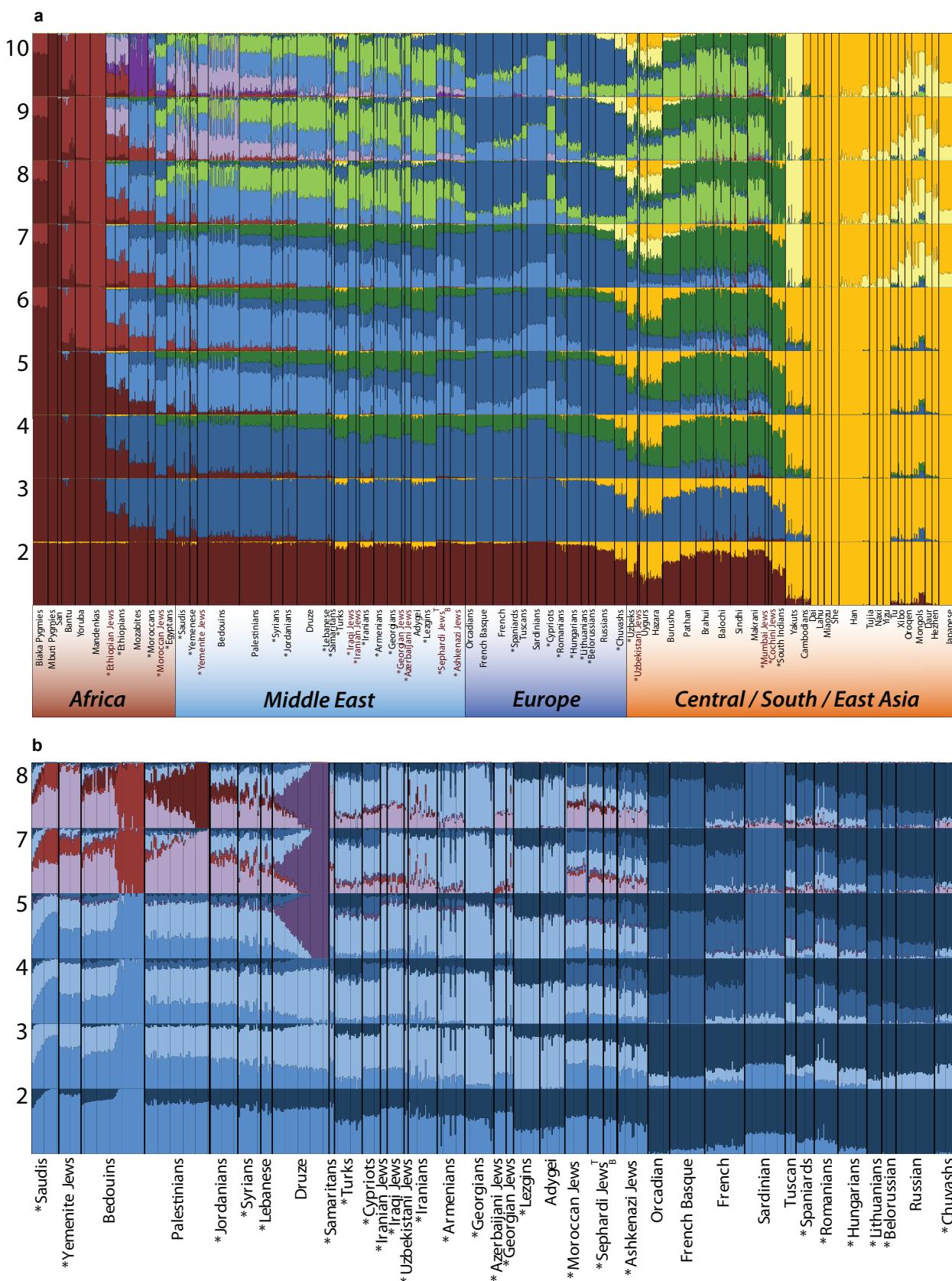
Supplementary Figure 2| Principal Component Analysis of the Old World High-Density Array Data.

a, Scatter plot of Old World individuals, showing the first two principal components. Here, the first PC (4.2% of variation, vertical axis) captures primarily differences between sub-Saharan Africans and the rest of the Old World. The second PC (3.4% of variation, horizontal axis) differentiates West Eurasians from South and East Asians. Axes of variation were scaled according to eigenvalues. Each letter code (**Supplementary Table 1**) corresponds to one individual and the color indicates population origin. **b**, Scatter plot of Old World individuals, showing PC1 and PC3. **c**, Scatter plot of Old World individuals, showing PC1 and PC4. Note that eigenvalues for PC3 and PC4 are ~8 times smaller than for PC1 and 2.



Supplementary Fig. 3 | Principal Component Analysis of West Eurasian High-Density

Array Data. A scatter plot of West Eurasian individuals together with individuals from different Jewish communities is shown. To facilitate extraction of PCs capturing genetic variation specific to West Eurasia we excluded populations showing clear membership in non-West Eurasian genetic components (Yemenites, Saudi Arabians, Ethiopian Jews, and Indian Jews). Axes of variation were scaled according to respective eigenvalues. Each letter code (**Supplementary Table 1**) corresponds to one individual and the color indicates region of origin.

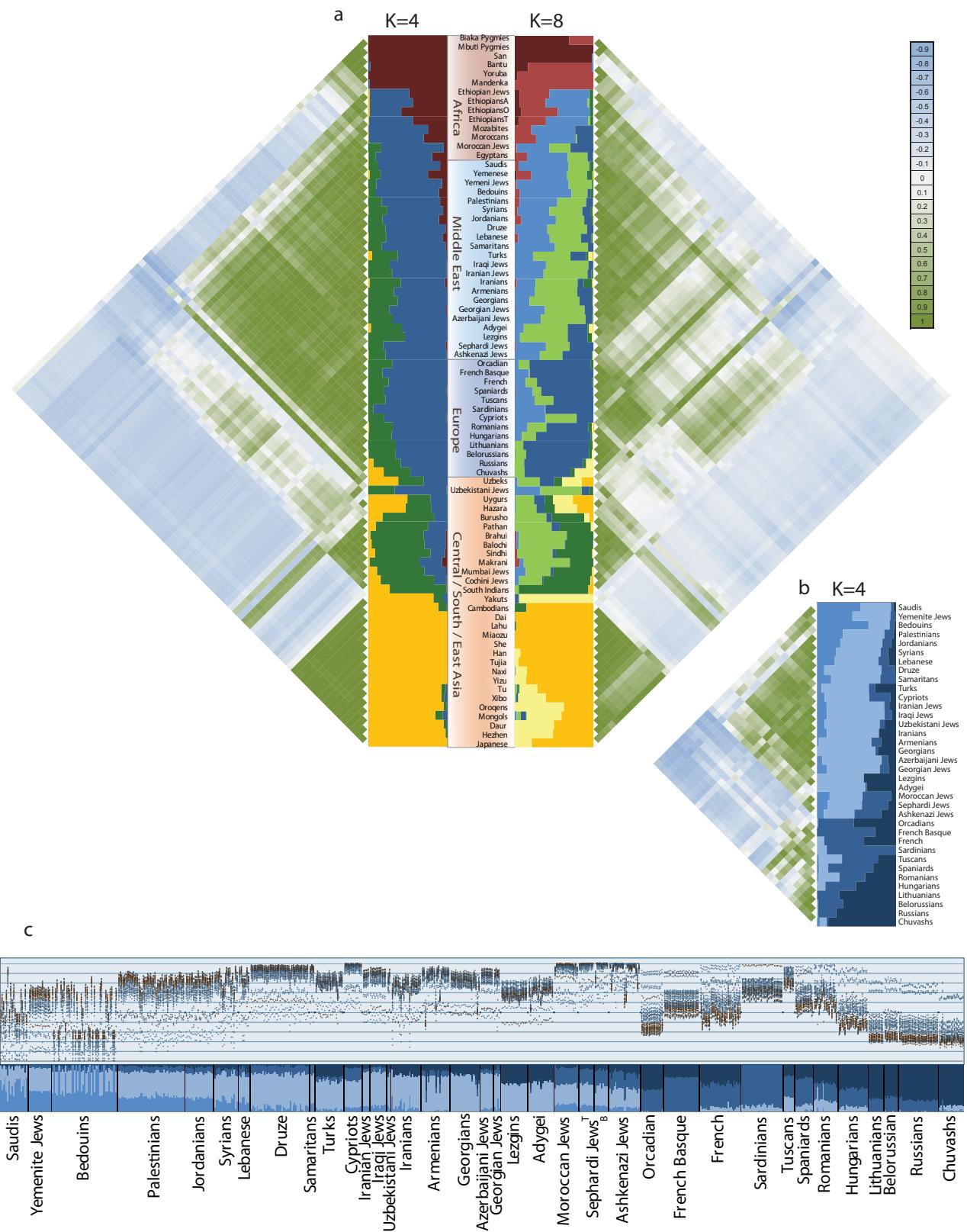


Supplementary Figure 4| ADMIXTURE analysis of the Old World and West Eurasian

samples at K=2 through K=10. a, Jewish communities are shown in colour and in bold. The letters T and B further specify Sephardi Jews from Turkey and Bulgaria, respectively.

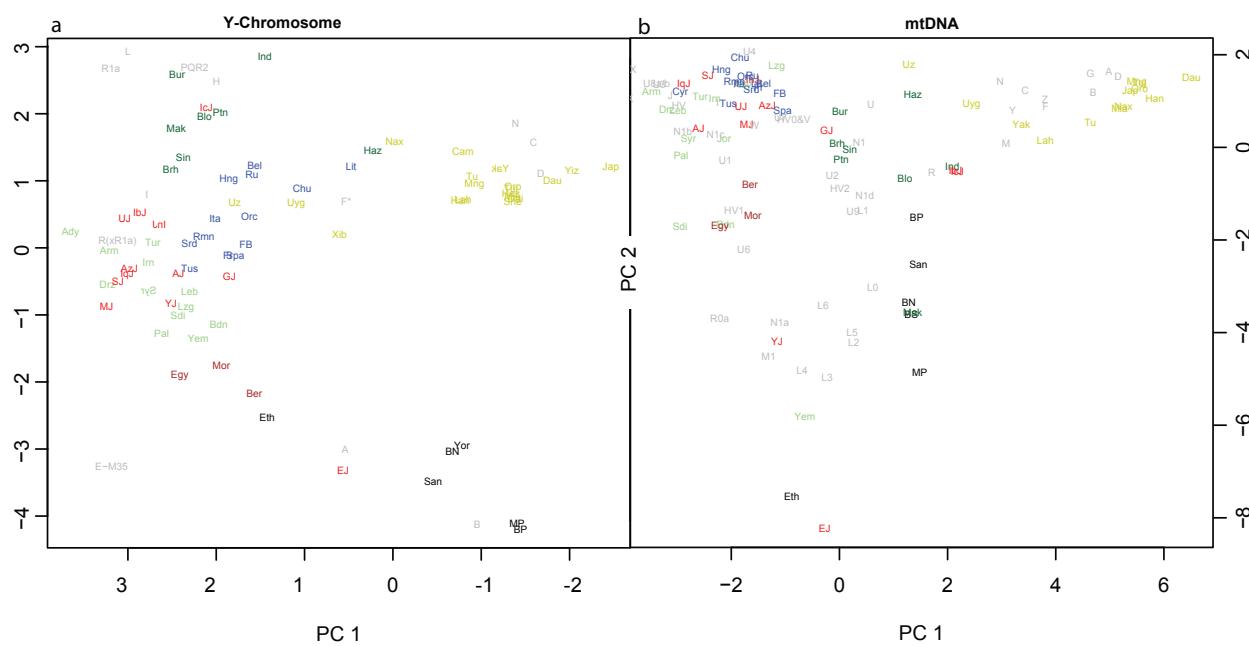
Populations introduced in this study for the first time and analysed together with the HGDP¹⁸

data are marked by an asterisk. See **Supplementary note 3 and 4** for more details.



Supplementary Figure 5 | Correlation coefficients. **a**, Individual inferred memberships in genetic components at K=4 and K=8 of the Old World dataset are averaged to the population

level. The resulting patterns are correlated and plotted in pair-wise correlation matrices. **b**, Equivalent to panel **a** using K=4 plot of the West Eurasian dataset (**Supplementary Note 3**). Colour bar depicts the colour-code of the correlation coefficients. **c**, Correlation coefficients of each Ashkenazi (blue dot) and Sephardi (brown dot) sample with every other individual are shown as columns of dots above the proportions plot (K=4 of the West Eurasian dataset) so that each column consists of 21 blue and 19 brown dots (i.e., corresponding to the number of Ashkenazi and Sephardi samples in the study) (**Supplementary Note 3**).



Supplementary Figure 6| Principal Component Analysis of Uniparental Data. Scatter plot of Jewish and non-Jewish populations based on haplogroup frequencies for **a**, Y-chromosome and **b**, mtDNA data. Abbreviations are as detailed in **Supplementary Table 1**. The populations, sample sizes, and haplogroups included in the analysis are summarized in **Supplementary Tables 4** and **5**. Haplogroups are also projected on the plot, using their normalized contribution to each principal component as scores.